

Halo-model signatures from 380 000 Sloan Digital Sky Survey luminous red galaxies with photometric redshifts

Chris Blake,^{1,2★} Adrian Collister³ and Ofer Lahav⁴

¹*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*

²*Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*

³*Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA*

⁴*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT*

Accepted 2007 April 26. Received 2007 April 25; in original form 2006 August 23

ABSTRACT

We analyse the small-scale clustering in ‘MegaZ-LRG’, a large photometric redshift catalogue of luminous red galaxies extracted from the imaging data set of the Sloan Digital Sky Survey. MegaZ-LRG, presented in a companion paper, spans the redshift range $0.4 < z < 0.7$ with an rms redshift error $\sigma_z \approx 0.03(1 + z)$, covering 5914 deg^2 to map out a total cosmic volume $2.5 h^{-3} \text{ Gpc}^3$. In this study we use 380 000 photometric redshifts to measure significant deviations from the canonical power-law fit to the angular correlation function in a series of narrow redshift slices, in which we construct volume-limited samples. These deviations are direct signatures of the manner in which these galaxies populate the underlying network of dark matter haloes. We cleanly delineate the separate contributions of the ‘one-halo’ and ‘two-halo’ clustering terms and fit our measurements by parametrizing the halo occupation distribution $N(M)$ of the galaxies. Our results are successfully fitted by a ‘central’ galaxy contribution with a ‘soft’ transition from zero to one galaxy, combined with a power-law ‘satellite’ galaxy component, the slope of which is a strong function of galaxy luminosity. The large majority of galaxies are classified as central objects of their host dark matter haloes rather than satellites in more massive systems. The effective halo mass of MegaZ-LRG galaxies lies in the range $\log_{10}(M_{\text{eff}}/h^{-1} M_{\odot}) = 13.61\text{--}13.80$ (increasing with redshift assuming large-scale normalization $\sigma_8 = 0.8$) for corresponding number densities in the range $n_g = 5.03\text{--}0.56 \times 10^4 h^{-3} \text{ Mpc}^{-3}$. Our results confirm the usefulness of the halo model for gaining physical insight into the patterns of galaxy clustering.

Key words: surveys – galaxies: haloes – large-scale structure of Universe.

1 INTRODUCTION

Photometric redshift surveys offer a route to delineating the large-scale structure of the Universe that is increasingly competitive with spectroscopic redshift surveys (Budavari et al. 2003; Seo & Eisenstein 2003; Amendola, Quercellini & Giallongo 2005; Dolney, Jain & Takada 2004; Blake & Bridle 2005; Phleps et al. 2006; Zhan et al. 2006; Blake et al. 2007; Padmanabhan et al. 2007). The ease with which modern imaging surveys can map large areas of sky to faint magnitude limits compensates for the absence of precise (but time-consuming) spectroscopic redshift measurements for individual galaxies. An absolute prerequisite, however, is the availability of high-quality photometric galaxy redshifts with known error distributions (established e.g. via spectroscopy of subsamples), together with accurate survey photometric calibration over large an-

gles of sky. Recent observational efforts have enabled both of these criteria to be satisfied.

The Sloan Digital Sky Survey (SDSS; York et al. 2000) has now provided an accurately calibrated imaging data set over roughly a fifth of the sky, which can be used to extract samples of galaxies in a uniform manner. In particular, a photometric catalogue of luminous red galaxies (LRGs) can be readily extracted using a series of well-understood colour and magnitude cuts (Eisenstein et al. 2001). Owing to their high luminosity and typical residence in the most massive dark matter haloes, LRGs are efficient tracers of cosmic structure across large volumes (Brown et al. 2003; Zehavi et al. 2005a; Eisenstein et al. 2005b). Moreover, these objects provide particularly reliable photometric redshifts owing to the strong spectral break at $\approx 4000 \text{ \AA}$ in the galaxy rest frame and the consequent rapid variation with redshift of their observed colours in the SDSS filter system (Padmanabhan et al. 2005; Collister et al. 2007). Furthermore, the photometric redshift error distribution of the LRGs can be accurately calibrated owing to the existence of spectroscopic observations of a subsample as part of the 2dF-SDSS LRG

★E-mail: cblake@astro.swin.edu.au

and Quasar (2SLAQ) survey at the Anglo-Australian Telescope (Cannon et al. 2006).

The combination of these data sets has allowed us to construct a large catalogue of more than 10^6 LRGs spanning the redshift range $0.4 < z < 0.7$ with an rms photometric redshift error $\sigma_z \equiv \sqrt{\langle(\delta z)^2\rangle} - \langle\delta z\rangle^2 \approx 0.03(1+z)$ where $\delta z \equiv z_{\text{phot}} - z_{\text{spec}}$. We have dubbed this catalogue ‘MegaZ-LRG’ (Collister et al. 2007). The data base covers almost 6000 deg^2 , or an effective volume $\approx 2.5 h^{-3} \text{ Gpc}^3$. We have already used this catalogue to measure the large-scale clustering of the galaxies on linear and quasi-linear scales via a power spectrum analysis and thereby extract measurements of cosmological parameters (Blake et al. 2007; see also Padmanabhan et al. 2007). In this study we turn to the small-scale clustering properties of the LRGs. Our goal is to present new measurements of the small-scale correlation function of LRGs at $z \sim 0.5$, and to connect these measurements to the physical manner in which these LRGs populate dark matter haloes.

It has long been known that different classes of galaxy possess different clustering properties, in a manner connected to their small-scale environments (for recent observational studies we refer the reader to Norberg et al. (2002); Budavari et al. (2003); Hogg et al. (2003); Zehavi et al. (2005b), and references therein). For many years these differing clustering properties could be adequately described by fitting a simple power-law function to the two-point galaxy correlation function on small scales (Peebles 1980). However, recent surveys have measured the clustering pattern accurately enough to detect deviations from the canonical clustering power law (e.g. Hawkins et al. 2003; Zehavi et al. 2004; Zheng 2004; Eisenstein et al. 2005a; Zehavi et al. 2005a,b; Phleps et al. 2006). These deviations provide an important insight into the processes of galaxy formation.

The richer structure in the galaxy clustering pattern revealed by recent surveys has been successfully interpreted in terms of the ‘halo model’ (e.g. Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Cooray & Sheth 2002; Kravtsov et al. 2004; Zehavi et al. 2004; Zheng 2004; Collister & Lahav 2005; Tinker et al. 2005; Zehavi et al. 2005b; Zheng et al. 2005). In this model, the small-scale clustering of a distribution of galaxies is linked to the underlying network of dark matter haloes, whose properties can be measured using cosmological simulations. A class of galaxies are assumed to populate haloes in accordance with a statistical ‘halo occupation distribution (HOD)’ as a function of the halo mass. The clustering then naturally separates into two components: the distribution of galaxies within individual haloes, which dominates on small scales ($\lesssim 1 \text{ Mpc}$), and the mutual clustering of galaxies inhabiting separate haloes, which dominates on larger scales ($\gtrsim 1 \text{ Mpc}$). The combination of these two terms can accurately model the observed scale-dependent features in the small-scale clustering pattern.

The outline of this paper is as follows. In Section 2, we briefly describe the ‘MegaZ-LRG’ data set. We then present measurements of the angular correlation function in narrow redshift slices together with simple power-law fits in Section 3. In Section 4, we introduce the framework of the halo model and in Section 5 we refit the clustering measurements by parametrizing the HOD of the LRGs, comparing our results to previous work in Section 6. Section 7 investigates a range of potential systematic photometric errors in the catalogue that may bias our results. We conclude in Section 8.

Throughout this study, we assume a fixed set of large-scale cosmological parameters: fractional matter density $\Omega_m = 0.3$, cosmological constant $\Omega_\Lambda = 0.7$, curvature $\Omega_k = 0$, Hubble parameter $h = 0.7$, fractional baryon density $f_b = \Omega_b/\Omega_m = 0.15$, slope of the

primordial power spectrum $n_s = 1$, and overall normalization of the power spectrum $\sigma_8 = 0.8$. These values are consistent with fits to the large-scale power spectrum of the LRGs (Blake et al. 2007) and to the latest measurements of the anisotropy spectrum of the cosmic microwave background (CMB) (Spergel et al. 2007). We note that there is some degeneracy between the assumed cosmological parameters and the fitted halo-model parameters, and we investigate how the best-fitting halo-model parameters depend on the values of σ_8 and n_s .

2 THE DATA SET

We analyse angular clustering in the ‘MegaZ-LRG’ galaxy data base, a photometric redshift catalogue of LRGs based on the imaging data set of the SDSS fourth Data Release. The construction of this catalogue is described in detail by Collister et al. (2007), and cosmological-parameter fitting to the angular power spectrum is presented by Blake et al. (2007). We only provide a brief description of the catalogue here, referring the reader to these two papers for more information.

MegaZ-LRG contains over 10^6 LRGs spanning the redshift range $0.4 < z < 0.7$ with an rms redshift error $\delta z \approx 0.03(1+z)$. The angular selection function is described by Blake et al. (2007) and encompasses 5914 deg^2 (the three southern SDSS stripes are excluded). The sample was selected from the SDSS imaging data base using a series of colour and magnitude cuts (Eisenstein et al. 2001; Collister et al. 2007), which amount to a magnitude-limited sample of LRGs with i -band magnitudes $17.5 < i < 20$. Photometric redshifts were derived for these galaxies using an Artificial Neural Network method, ANNz (Firth, Lahav & Somerville 2003; Collister & Lahav 2004), constrained by a spectroscopic subsample of $\approx 13\,000$ galaxies obtained by the 2SLAQ survey (Cannon et al. 2006). In this paper, we analyse a ‘conservative’ version of the data base in which the selection cuts applied to the imaging data base are identical to those used to produce the great majority of the spectroscopic subsample (e.g. the faint magnitude limit is brightened to $i = 19.8$). Star–galaxy separation cuts were applied both in the initial selection from the SDSS data base and via the neural network analysis. These cuts produced a catalogue with 644 903 entries, amongst which there is a 1.5 per cent M-star contamination (see Blake et al. 2007).

We used the photometric redshifts to divide the sample into four narrow redshift slices of width $\Delta z_{\text{phot}} = 0.05$ between $z = 0.45$ and 0.65 (there are very few galaxies in the catalogue outside this redshift range). We then applied a luminosity threshold in each redshift range to create a ‘volume-limited’ sample of galaxies, as assumed by the halo model. The luminosity threshold is given by the faint apparent magnitude limit $i = 19.8$ applied at the most-distant redshift of each slice. We calculated luminosities for each galaxy using LRG K -corrections from the 2SLAQ survey (Wake et al. 2006), not including an evolution correction. The faint absolute i -band magnitude limits for each redshift slice are then $M_i - 5 \log_{10} h = (-22.23, -22.56, -22.87, -23.20)$. The number of galaxies remaining in each redshift slice after the luminosity cut was $N = (168287, 118863, 70229, 27203)$ with corresponding surface densities $(28.5, 20.1, 11.9, 4.6) \text{ deg}^{-2}$. The spectroscopic redshift distribution of galaxies in each photo- z slice (including the luminosity cut) can be deduced using the 2SLAQ spectroscopic subsample. As shown by Blake et al. (2007), the spec- z probability distribution for each slice is well described by a Gaussian function; the mean μ and standard deviation σ for each photo- z slice are listed in Table 1 (note that these values are slightly different from those

Table 1. Power-law fits to the angular correlation function in the four redshift slices. Column 2 records the surface density of the galaxy sample in each redshift slice. Column 3 is the threshold galaxy absolute i -band magnitude in each slice for the volume-limited sample. Columns 4 and 5 list the parameters of the Gaussian redshift distribution for each slice, $p(z) \propto \exp[-(z - \mu)^2/2\sigma^2]$. Columns 6 and 7 show the best-fitting power-law parameters $w(\theta) = a\theta^{1-\gamma}$, where θ is in degrees. The associated errors for a and γ are obtained by marginalizing over values of the other parameter. Note that the amplitudes have been corrected upwards by a stellar contamination factor $(1 - f)^{-2} = 1.03$. Column 8 is the corresponding χ^2 -statistic χ^2_{pl} evaluated using the full covariance matrix (which we compare with 25 degrees of freedom). Column 9 lists the inferred galaxy clustering length r_0 of the spatial correlation function $\xi(r) = (r/r_0)^{-\gamma}$.

Redshift slice	Density (deg ⁻²)	$M_i - 5 \log_{10} h$	μ	σ	$10^2 \times a$ (deg ^{$\gamma-1$})	γ	χ^2_{pl}	r_0 (h^{-1} Mpc)
$0.45 < z < 0.5$	28.5	-22.23	0.476	0.034	2.84 ± 0.12	1.96 ± 0.01	177.1	7.2 ± 0.2
$0.5 < z < 0.55$	20.1	-22.56	0.528	0.040	2.79 ± 0.14	1.96 ± 0.01	132.8	8.0 ± 0.2
$0.55 < z < 0.6$	11.9	-22.87	0.574	0.043	2.93 ± 0.18	1.94 ± 0.02	71.8	8.5 ± 0.3
$0.6 < z < 0.65$	4.6	-23.20	0.629	0.052	2.78 ± 0.19	1.94 ± 0.02	34.9	9.3 ± 0.3

listed in Blake et al. 2007, due to the additional luminosity threshold applied in this study). These redshift distributions are used to project the model correlation function to fit the observed angular clustering in each slice.

3 ANGULAR CORRELATION FUNCTION MEASUREMENTS

3.1 Method

We used the Landy & Szalay (1993) estimator to measure $w(\theta)$ for each photometric redshift slice in 30 logarithmically spaced angular separation bins between $\theta = 0.001$ and 1° . For each redshift slice we generated 10 random data sets across the survey geometry, each containing the same number of galaxies as the survey data sets. We assume that the 1.5 per cent stellar contamination is distributed evenly across the redshift slices in an unclustered fashion such that the amplitude of the measured angular correlation function is simply reduced by a constant factor $(1 - f)^2$ where $f = 0.015$. We corrected our estimates of $w(\theta)$ and corresponding errors upwards by this factor (of 3.0 per cent).

3.2 Error determination

We estimated the covariance matrix of the errors in the separation bins using the technique of jack-knife resampling. For each measurement of $w(\theta)$ we divided the survey area into $N = 393$ subfields of constant area (of 13.3 deg^2) using a grid of right ascension and declination divisions (see Fig. 1). When creating the grid, we allowed a subfield to contain up to 20 per cent fractional area beyond survey edges or of survey holes. The number of subfields was chosen to be sufficient for estimating each unique element of the covariance matrix used in the model fitting with statistical independence. However, we checked that our best-fitting parameters did not depend on the number of subfields, or on the restriction of our fits to the most significant principal components of the covariance matrix using singular value decomposition techniques.

Having defined the grid of N subfields, we then measured $w(\theta)$ N times, which we label as $w_i(\theta)$ from $i = 1$ to N , in each case omitting just one of the subfields (and using the remaining $N - 1$ fields). The covariance of the measurements between separation bins j and k was deduced as

$$C_{jk} \equiv \langle w(\theta_j) w(\theta_k) \rangle - \langle w(\theta_j) \rangle \langle w(\theta_k) \rangle \quad (1)$$

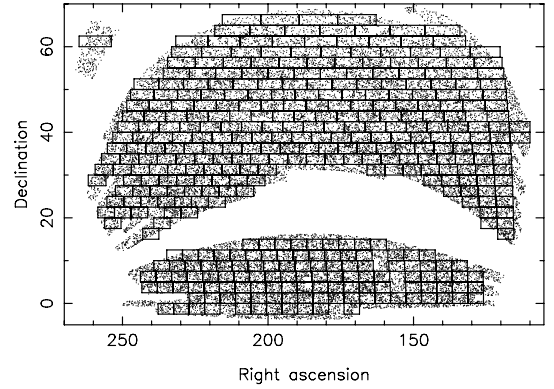


Figure 1. The grid of $N = 393$ subfields of equal area, defined by constant right ascension and declination boundaries, which was used to estimate the error in the angular correlation function via jack-knife resampling of the measurement.

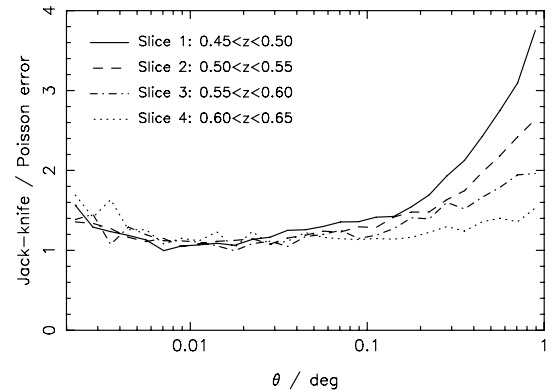


Figure 2. The error in the angular correlation function determined by jack-knife resampling for the four redshift slices, normalized by the result assuming simple Poisson statistics.

$$\approx (N - 1) \left[\frac{\sum_i^N w_i(\theta_j) w_i(\theta_k)}{N} - \overline{w(\theta_j)} \times \overline{w(\theta_k)} \right], \quad (2)$$

where $\overline{w(\theta_j)} = \sum_i^N w_i(\theta_j)/N$.

Fig. 2 plots the jack-knife errors in the separation bins for each redshift slice, normalized by the error determined assuming simple Poisson statistics (for which the error in the data pair count of DD objects in a separation bin is \sqrt{DD}). For most of the angular range

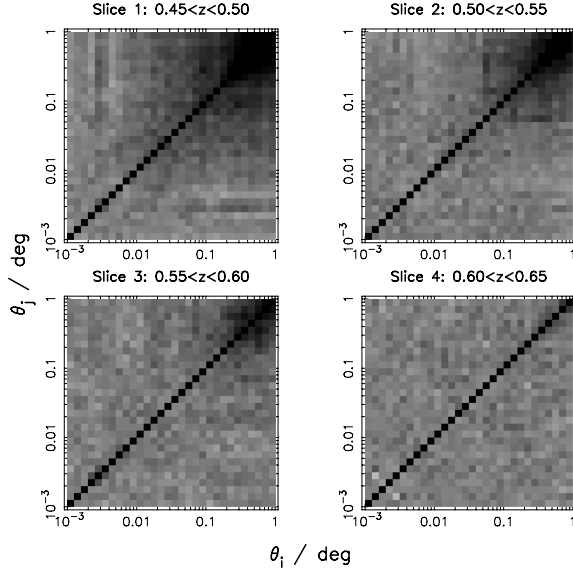


Figure 3. A grey-scale plot of the correlation coefficient r of equation (3), indicating the degree of covariance between different separation bins for each redshift slice. The correspondence of r to the grey-scale is chosen as $r = -0.5$ (white) to $r = 0.5$ (black).

under investigation, the jack-knife errors are less than 50 per cent higher than those predicted by Poisson statistics. For the largest angular scales θ considered, the jack-knife error increases relative to the Poisson error, owing to the increasing importance of edge effects and the heightened ‘cosmic variance’ owing to the reduced number of independent cells of size θ that can be accommodated by the data set. This is the familiar result from clustering measurements that Poisson noise dominates on small scales, and cosmic variance dominates on large scales. The full covariance matrices are displayed in Fig. 3 by plotting in grey-scale the ‘correlation coefficient’ between

two separation bins i and j :

$$r(i, j) = \frac{\text{Cov}(i, j)}{\sqrt{\text{Cov}(i, i)\text{Cov}(j, j)}}. \quad (3)$$

Angular correlation function measurements in large-separation bins are positively correlated. The amplitude of these correlations decreases with redshift because, as the number density of the sample reduces, Poisson noise becomes more important compared to cosmic variance. These covariance matrices are always used in our model fitting.

3.3 Power-law fits

As an initial step we fitted power laws $w(\theta) = a\theta^{1-\gamma}$ to the measured angular correlation functions in the four redshift slices (using χ^2 -minimization with the full covariance matrix). We excluded the first three data points with separations $\theta < 0.002$ ($= 7$ arcsec), as these scales are potentially affected by astronomical seeing and issues of galaxy merging and blending (the decrease in the value of the correlation functions at the smallest scales in Fig. 4 is not a property of galaxy clustering). The best-fitting power-law parameters (a, γ) are listed in Table 1 and the models and data points are displayed in Fig. 4. The best-fitting slopes γ for the redshift slices take values in the range 1.94–1.96, consistent with previous studies of LRGs (e.g. Eisenstein et al. 2005a; Zehavi et al. 2005a). However, as indicated by the high values of the minimum χ^2 -statistic χ^2_{pl} in Table 1 compared to the 25 degrees of freedom, a power law is not a good fit to the data. The fact that we detect deviations from simple power-law clustering with high significance indicates that our photo- z survey can be used to fit more complex and physically-insightful models to the small-scale clustering pattern.

Nevertheless, we can use the power-law amplitude of the angular clustering to estimate the three-dimensional clustering length r_0 of the population of galaxies, using the known redshift distribution $p(z)$ in each slice. If we define the spatial two-point correlation

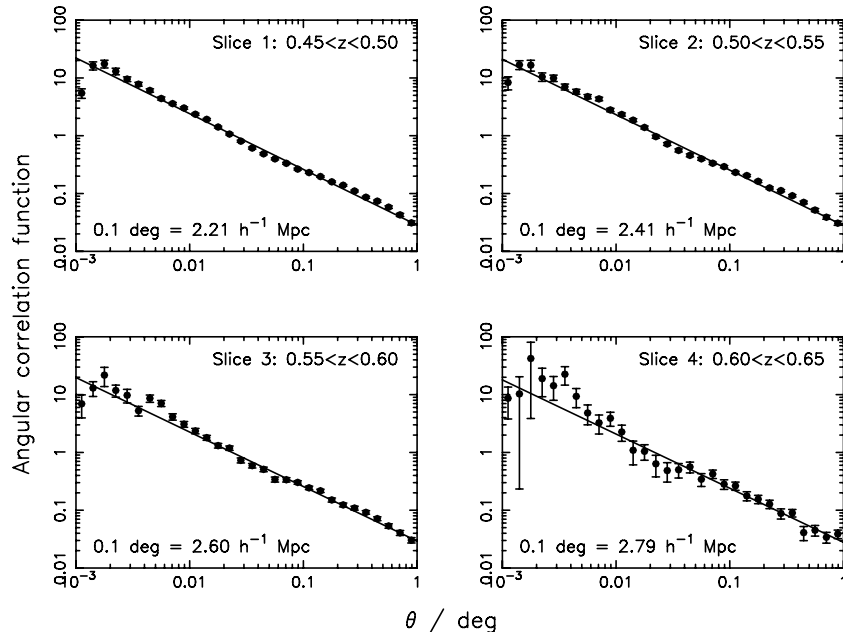


Figure 4. The measured angular correlation function in the four redshift slices together with the best-fitting power laws. The turnover of the correlation function towards very small angular separations is not a property of galaxy clustering; we restrict our fit to separations $\theta > 0.002$. The comoving spatial scale corresponding to an angular separation 0.1 at the median redshift of each slice is displayed in the bottom left-hand corner of each plot.

function $\xi(r) = (r/r_0)^{-\gamma}$, then the amplitude of the power-law angular correlation function $w(\theta) = a\theta^{1-\gamma}$ follows from Limber's equation:

$$a = C_\gamma r_0^\gamma \int dz p(z)^2 \left(\frac{dx}{dz} \right)^{-1} x(z)^{1-\gamma}, \quad (4)$$

where $x(z)$ is the comoving radial coordinate at redshift z and $C_\gamma = \Gamma(\frac{1}{2})\Gamma(\frac{\gamma}{2} - \frac{1}{2})/\Gamma(\frac{\gamma}{2})$. In Table 1, we determine the corresponding values of r_0 and γ for each redshift slice. The amplitude of the clustering length, $r_0 = 7.2\text{--}9.3 h^{-1}$ Mpc, is consistent with highly biased massive galaxies. Our small-scale angular clustering measurements are not affected by redshift-space distortions, since the photo- z errors are much bigger than the peculiar velocities of the galaxies. Correlated redshift-space distortions are important on larger scales, however, as discussed in Blake et al. (2007) and Padmanabhan et al. (2007).

The clustering amplitude systematically increases with redshift for the following two reasons.

(i) Each volume-limited sample in successive redshift slices has the same limiting apparent magnitude hence higher luminosity threshold. The higher-redshift galaxies are hence preferentially more luminous and more strongly clustered (e.g. Norberg et al. 2002; Zehavi et al. 2005b). In Table 1, we list the threshold absolute i -band magnitude M_i of galaxies in each redshift slice, calculated using an LRG K -correction (Wake et al. 2006).

(ii) In standard models of the evolution of galaxy clustering, the bias factor of a class of galaxies increases with redshift in opposition to the decreasing linear growth factor, in order to reproduce the observed approximate constancy of the small-scale comoving clustering length (e.g. Lahav et al. 2002). Simple models for this effect such as $b(z) = 1 + (b_0 - 1)/D(z)$ (Fry 1996) or $b(z) = b_0/D(z)$, where $D(z)$ is the linear growth factor, predict an evolution in bias across our analysed redshift range of $\Delta b \approx 0.2$.

These trends are in good agreement with our measurements of the amplitude of the large-scale clustering pattern (Blake et al. 2007).

4 HALO-MODEL FRAMEWORK

We use the halo model of galaxy clustering to produce model spatial correlation functions $\xi(r)$ to fit to our measurements. We summarize the ingredients of our model here. Further details can be found in, for example, Seljak (2000), Cooray & Sheth (2002), Berlind & Weinberg (2002), Kravtsov et al. (2004), Zehavi et al. (2004), Zheng (2004), Zehavi et al. (2005b) and Tinker et al. (2005).

In the halo-model framework, the clustering functions are expressed as a sum of components due to pairs of galaxies within a single dark matter halo (the ‘one-halo term’ ξ_1) and to pairs of galaxies inhabiting separate haloes (the ‘two-halo term’ ξ_2):

$$\xi(r) = [1 + \xi_1(r)] + \xi_2(r) \quad (5)$$

where the ‘1+’ at the start of the expression arises because the total number of galaxy pairs ($\propto 1 + \xi$) is the sum of the number of pairs from single haloes ($\propto 1 + \xi_1$) and from different haloes ($\propto 1 + \xi_2$). The two terms dominate on different scales, with the one-halo term only important on small scales $\lesssim 1$ Mpc.

The fundamental ingredient of the galaxy halo model is the *HOD*, which describes the probability distribution for the number of galaxies N hosted by a dark matter halo as a function of its mass M . In order to construct the one-halo and two-halo two-point clustering terms, we require the first and second factorial moments of the HOD, $\langle N | M \rangle$ and $\langle N(N - 1) | M \rangle$. We make the assumption that the first

galaxy to be hosted by a halo lies at the centre of the halo, and any remaining galaxies are classified as ‘satellites’ and distributed in proportion to the halo mass profile. We apply different HODs for the central and satellite galaxies, $\langle N_c | M \rangle$ and $\langle N_s | M \rangle$, respectively, where

$$\langle N | M \rangle = \langle N_c | M \rangle (1 + \langle N_s | M \rangle) \quad (6)$$

Equation (6) takes its form because a halo can only host a satellite galaxy if it already contains a central galaxy. We will use the notation $N_c(M) = \langle N_c | M \rangle$, $N_s(M) = \langle N_s | M \rangle$ and $N(M) = \langle N | M \rangle$ in the equations that follow.

4.1 The one-halo term $\xi_1(r)$

The one-halo galaxy correlation function is composed of contributions from central–satellite pairs and satellite–satellite pairs. It is convenient to evaluate these two contributions separately. The one-halo correlation function for central–satellite pairs is given by

$$1 + \xi_{1,c-s}(r) = \int_{M_{\text{vir}}(r)}^\infty dM n(M) \frac{N_c(M) N_s(M)}{n_g^2/2} \frac{\rho(r|M)}{M}, \quad (7)$$

where n_g is the galaxy number density, $n(M)$ is the halo mass function, and $\rho(r|M)$ is the halo density profile. The lower limit for the integral is the halo mass M corresponding to a virial radius r , given that less-massive haloes have smaller radii and cannot contribute any central–satellite galaxy pairs with a separation r :

$$M_{\text{vir}}(r) = \frac{4}{3} \pi r^3 \bar{\rho} \Delta, \quad (8)$$

where $\bar{\rho} = 2.78 \times 10^{11} \Omega_m h^2 \text{M}_\odot \text{Mpc}^{-3}$ is the co-moving background density of the Universe, and $\Delta = 200$ is the critical overdensity for virialization.

It is simplest to evaluate the one-halo correlation function for satellite–satellite pairs in a Fourier space (where convolutions become multiplications). The power spectrum is

$$P_{1,s-s}(k) = \int_0^\infty dM n(M) \frac{N_c(M) N_s^2(M)}{n_g^2} |u(k|M)|^2, \quad (9)$$

where $u(k|M)$ is the Fourier transform of the halo density profile $\rho(r|M)$. Because satellite galaxies are Poisson-distributed, we can write $\langle N_s(N_s - 1) \rangle = \langle N_s \rangle^2$ to obtain the above equation. The correlation function corresponding to equation (9) is then

$$\xi_{1,s-s}(r) = \frac{1}{2\pi^2} \int_0^\infty dk P_{1,s-s}(k) k^2 \frac{\sin kr}{kr}. \quad (10)$$

The total one-halo correlation function is then derived as

$$\xi_1 = \xi_{1,c-s} + \xi_{1,s-s} \quad (11)$$

4.2 The two-halo term $\xi_2(r)$

The two-halo galaxy correlation function at a separation r is evaluated from the scale-dependent two-halo power spectrum $P_2(k, r)$:

$$P_2(k, r) = P_m(k) \times \left[\int_0^{M_{\text{lim}}(r)} dM n(M) b(M, r) \frac{N(M)}{n'_g(r)} u(k|M) \right]^2, \quad (12)$$

where $P_m(k)$ is the non-linear matter power spectrum at the survey redshift, $b(M, r)$ is the scale-dependent halo bias at separation r , and $n'_g(r)$ is the restricted galaxy number density at a separation r :

$$n'_g(r) = \int_0^{M_{\text{lim}}(r)} dM n(M) N(M). \quad (13)$$

The mass truncation $M_{\text{lim}}(r)$ must be included to incorporate the effects of halo exclusion: more massive haloes would overlap at separation r . We derive the mass limit using the ‘ n'_g -matched’ approximation described in Tinker et al. (2005). First, we calculate the restricted number density using equation (B13) in Tinker et al., which includes the effects of triaxiality:

$$n'_g(r) = \int_0^\infty dM_1 n(M_1) N(M_1) \times \int_0^\infty dM_2 n(M_2) N(M_2) P(r, M_1, M_2), \quad (14)$$

where $P(r, M_1, M_2)$ quantifies the probability of non-overlapping haloes of masses M_1 and M_2 with separation r . Defining $x = r/(R_1 + R_2)$, where R_1 and R_2 are the virial radii corresponding to masses M_1 and M_2 , and using $y = (x - 0.8/0.29)$, then Tinker et al. obtain $P(y) = 3y^2 - 2y^3$ from simulations. Given this value of $n'_g(r)$, we increase the value of $M_{\text{lim}}(r)$ in equation (13) to produce a matching number density. This value is then used in equation (12) to produce the two-halo power spectrum $P_2(k, r)$.

Following Tinker et al. (2005), we assumed the following model for the scale-dependent bias:

$$b^2(M, r) = b^2(M) \frac{[1 + 1.17 \xi_m(r)]^{1.49}}{[1 + 0.69 \xi_m(r)]^{2.09}}, \quad (15)$$

where $\xi_m(r)$ is the non-linear matter correlation function. We derived matter power spectra using the ‘CAMB’ software package (Lewis, Challinor & Lasenby 2000), including corrections for non-linear growth of structure using the fitting formulae of Smith et al. (2003) (‘halofit=1’ in CAMB). We output power spectra at the mean redshift of each slice, and obtained the correlation function using a Fourier transform. The two-halo galaxy correlation function is obtained via the Fourier transform of equation (12):

$$\xi'_2(r) = \frac{1}{2\pi^2} \int_0^\infty dk P_2(k, r) k^2 \frac{\sin kr}{kr}, \quad (16)$$

with the number of galaxy pairs corrected from the restricted galaxy density to the entire galaxy population:

$$1 + \xi_2(r) = \left[\frac{n'_g(r)}{n_g} \right]^2 [1 + \xi'_2(r)]. \quad (17)$$

4.3 Halo mass and bias functions

The *halo mass function*, $n(M)$, describes the number density of haloes as a function of mass M . We introduce the new mass variable ν following Press & Schechter (1974):

$$n(M) dM = \frac{\bar{\rho}}{M} f(\nu) d\nu. \quad (18)$$

The new mass variable ν is defined by $\nu = [\delta_{\text{sc}}/\sigma(M, z)]^2$, where δ_{sc} is the linear-theory prediction for the present-day overdensity of a region which would undergo spherical collapse at redshift z , and $\sigma^2(M, z)$ is the variance of the linear power spectrum in a spherical top hat which contains average mass M :

$$\sigma^2(M, z) = [D(z)^2/2\pi^2] \int_0^\infty dk k^2 P_{\text{lin}}(k) W^2(kR), \quad (19)$$

where $W(x) = (3/x^3)[\sin x - x \cos x]$, $M = (4/3)\pi R^3 \rho$, $D(z)$ is the linear growth factor at redshift z , and $P_{\text{lin}}(k)$ is the linear power

spectrum at redshift zero. We approximate $\delta_{\text{sc}} = 1.686$ independently of redshift. We use the Jenkins et al. (2001) model for the mass function:

$$\nu f(\nu) = (1/2)a_1 \exp[-|\ln(\sqrt{\nu/\delta_{\text{sc}}}) + a_2|^{a_3}] \quad (20)$$

where $a_1 = 0.315$, $a_2 = 0.61$ and $a_3 = 3.8$.

The *halo bias function* $b(M)$ describes the biasing of a halo of mass M with respect to the overall dark matter distribution. We use the Sheth, Mo & Tormen (2001) model for the bias function with the revised parameters stated in Tinker et al. (2005):

$$b(\nu) = 1 + \frac{1}{\delta_{\text{sc}}} \times \left\{ q\nu + s(q\nu)^{1-t} - \frac{q^{-1/2}}{1 + s(1-t)[1 - (t/2)](q\nu)^{-t}} \right\}, \quad (21)$$

where the constants $q = 0.707$, $s = 0.35$ and $t = 0.8$.

4.4 Halo profiles

We use the Navarro, Frenk & White (1997) dark matter halo density profile:

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \quad (r < r_{\text{vir}}), \quad (22)$$

where r_s is the characteristic scale radius and ρ_s provides the normalization. The profile is truncated at the virial radius r_{vir} , which is obtained from the halo mass via

$$r_{\text{vir}} = \left(\frac{3M}{4\pi\Delta\bar{\rho}} \right)^{1/3}. \quad (23)$$

We parametrize the profile in terms of the *concentration parameter* $c = r_{\text{vir}}/r_s$. The normalization for the mass M is

$$M = \int_0^{r_{\text{vir}}} \rho(r) 4\pi r^2 dr = 4\pi\rho_s r_s^3 \left[\ln(1+c) - \frac{c}{1+c} \right]. \quad (24)$$

We assume that the concentration parameter c depends on halo mass M in a manner calibrated by numerical simulations (Bullock et al. 2001; Zehavi et al. 2004):

$$c(M, z) = [11/(1+z)](M/M_0)^{-0.13} \quad (25)$$

where M_0 is obtained from equation (19) by setting $\sigma(M_0, 0) = \delta_{\text{sc}}$. For our adopted cosmological model we obtain $M_0 = 12.64 h^{-1} M_\odot$, resulting in a concentration $c = 5$ for a halo of mass $M = 10^{14} h^{-1} M_\odot$ at $z = 0.5$. This assumption is consistent with the measurement by Mandelbaum et al. (2006b) of the concentration parameter for LRGs using galaxy–galaxy weak lensing.

4.5 Halo occupation distribution

The HOD is a parametrized description of how galaxies populate dark matter haloes as a function of the halo mass M . As described above, we separate the distribution into separate HODs for central galaxies and for satellite galaxies. We adopt simple models motivated by results from simulations and semi-analytic calculations (e.g. Kauffmann, Nusser & Steinmetz 1997; Benson et al. 2000; Berlind et al. 2003; Kravtsov et al. 2004).

For central galaxies, our basic model is a step function such that haloes above a minimum mass threshold, M_{cut} , contain a single central galaxy and haloes below this threshold contain no galaxies. However, our observational data consist of galaxy luminosities (with

a scatter resulting from the photo- z errors) rather than galaxy masses. We therefore follow Zheng et al. (2005) and ‘soften’ the transition from zero to one galaxy with a further parameter σ_{cut} such that

$$\langle N_c | M \rangle = 0.5 \left[1 + \text{erf} \left(\frac{\log_{10}(M/M_{\text{cut}})}{\sigma_{\text{cut}}} \right) \right]. \quad (26)$$

The scatter in halo mass for a fixed galaxy luminosity results from galaxy formation physics and (in the case of this study) from photo- z errors.

For satellite galaxies, a power law in mass provides a good description for the mean occupation number in simulations:

$$\langle N_s | M \rangle = \left(\frac{M}{M_0} \right)^\beta. \quad (27)$$

Introducing a ‘cut-off’ to the satellite HOD of the form

$$\langle N_s | M \rangle = \left(\frac{M - M_1}{M_0} \right)^\beta \quad (28)$$

(Zheng et al. 2005) did not significantly improve the fit of the model to the data. Our halo-model power spectrum is hence specified by four variables: M_{cut} , σ_{cut} , M_0 and β .

In addition we must satisfy one extra constraint: the galaxy number density

$$n_g = \int_0^\infty dM n(M) N(M) \quad (29)$$

must match the observed number density in each redshift slice. We match this constraint by fixing the variable M_{cut} for each choice of σ_{cut} , M_0 and β . Our model hence contains three independent parameters. Fig. 5 displays an example halo-model correlation function and the component one-halo and two-halo terms.

Useful quantities which can be derived from the HOD are the effective large-scale bias

$$b_g = \int dM n(M) b(M) \frac{N(M)}{n_g}, \quad (30)$$

and the effective mass M_{eff} of the HOD

$$M_{\text{eff}} = \int dM M n(M) \frac{N(M)}{n_g}. \quad (31)$$

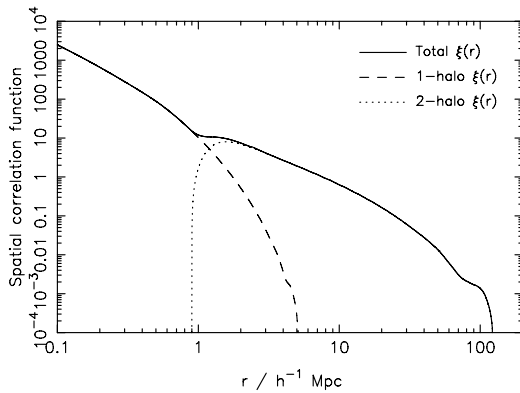


Figure 5. An example halo-model galaxy correlation function at $z = 0.5$ for parameters $\sigma_{\text{cut}} = 0.3$, $\log_{10}(M_0/h^{-1} M_\odot) = 14.3$ and $\beta = 1.5$. Matching galaxy density $n_g = 5 \times 10^{-4} h^3 \text{ Mpc}^{-3}$ requires $\log_{10}(M_{\text{cut}}/h^{-1} M_\odot) = 12.98$. The separate one-halo and two-halo components are shown, which cross over in dominance at a scale $\sim 1 h^{-1} \text{ Mpc}$. This model has a linear bias $b_g = 1.82$ and an effective halo mass $\log_{10}(M_{\text{eff}}/h^{-1} M_\odot) = 13.48$.

We can also determine the average fraction of central or satellite galaxies in the sample, for example, for central galaxies

$$f_c = \frac{\int dM n(M) N_c(M)}{\int dM n(M) (N_c(M) [1 + N_s(M)])} \quad (32)$$

and $f_s = 1 - f_c$ for satellite galaxies.

4.6 Conversion to angular correlation function

Knowing the redshift distribution $p(z)$ of a galaxy population, we can project the spatial galaxy correlation function $\xi(r)$ to an angular correlation function $w(\theta)$ using Limber’s equation:

$$w(\theta) = 2 \int_0^\infty dx f(x)^2 \int_0^\infty du \xi(r = \sqrt{u^2 + x^2 \theta^2}), \quad (33)$$

where $f(x)$ describes the radial distribution of sources as

$$f(x) = \frac{p(z)}{dx/dz(z)}, \quad (34)$$

where z is the redshift corresponding to comoving radial coordinate $x(z)$.

5 PARAMETER FITS

We fitted the three-parameter halo model (σ_{cut} , M_0 , β) to the observed angular correlation functions in each redshift slice, fixing the remaining parameter M_{cut} by matching to the observed galaxy number density n_g . We used a combination of a coarse grid-based search and a downhill-simplex method to locate the minimum value of the χ^2 -statistic, using the full covariance matrix. We then employed a fine grid-based method to explore the χ^2 -surface around the minimum and determined the errors in the fitted parameters (by marginalizing over the other model parameters). The mean and standard deviation of each model parameter, marginalizing over the other parameters, is listed in Table 2.

In Fig. 6, we plot the best-fitting halo-model correlation functions together with the data. We divide the results by the best-fitting power-law model from Section 3.3 for increased clarity. We note immediately that the halo-model framework has successfully reproduced the deviations from the power law, owing to the separate contributions of the one-halo and two-halo terms. The result is a good fit of model to data, with the addition of only one extra parameter compared to the original power-law fit. The minimum values of χ^2 are around 30 (for 24 degrees of freedom). Inspection of Fig. 6 reveals that the best-fitting model always describes the transition region between the one-halo and two-halo terms very well, and that the main source of discrepancy is at very small scales ($< 200 \text{ kpc}$) where the data points lie systematically above the model prediction.

In Table 2, we also list the derived values of the galaxy bias factor b_g and effective halo mass M_{eff} for each redshift slice (calculated using equations 30 and 31). The errors in these quantities are obtained by evaluating their values over the fine grid of halo-model parameters and weighting by the appropriate probability for the fit of model to data at each grid point. The range of linear bias, increasing with redshift from $b_g \approx 1.92$ (at $z = 0.475$) to $b_g \approx 2.62$ (at $z = 0.625$), agrees well with fits to the large-scale power spectrum of the LRGs (Blake et al. 2007), allowing for the differing values of the normalization σ_8 [in this study we have assumed $\sigma_8 = 0.8$, whereas the bias values in Blake et al. (2007) are quoted for $\sigma_8 = 1$]. The effective mass, ranging from $M_{\text{eff}} = 10^{13.61} - 10^{13.80} h^{-1} M_\odot$ over the same redshift range, confirms that LRGs are hosted by massive

Table 2. Halo-model fits to the angular correlation function in the four redshift slices. Column 2 records the observed galaxy number density which our models are constrained to match. Columns 3, 4 and 5 list the best-fitting values and errors of the halo-model parameters σ_{cut} , M_0 and β defined by equations (25) and (26). Column 6 displays the inferred value of the final halo-model parameter M_{cut} . In Columns 7 and 8, we evaluate the corresponding galaxy bias factor b_g (equation 30) and effective halo mass M_{eff} (equation 31). Column 9 records the minimum value of the χ^2 -statistic for the halo model, χ^2_{halo} , evaluated using the full covariance matrix (which we compare with 24 degrees of freedom). Base-10 logarithms are used in this table.

Redshift slice	$10^4 \times n_g$ ($h^3 \text{ Mpc}^{-3}$)	σ_{cut}	$\log(M_0/M_\odot h^{-1})$	β	$\log(M_{\text{cut}}/M_\odot h^{-1})$	b_g	$\log(M_{\text{eff}}/M_\odot h^{-1})$	χ^2_{halo}
$0.45 < z < 0.5$	5.03	0.21 ± 0.11	14.09 ± 0.01	1.57 ± 0.02	12.98	1.92 ± 0.02	13.61 ± 0.01	22.1
$0.5 < z < 0.55$	3.07	0.07 ± 0.07	14.22 ± 0.01	1.69 ± 0.04	13.12	2.15 ± 0.01	13.67 ± 0.01	28.3
$0.55 < z < 0.6$	1.60	0.24 ± 0.12	14.39 ± 0.01	1.87 ± 0.07	13.35	2.38 ± 0.04	13.74 ± 0.01	36.8
$0.6 < z < 0.65$	0.56	0.53 ± 0.14	14.76 ± 0.10	1.80 ± 0.36	13.79	2.62 ± 0.09	13.80 ± 0.03	40.1

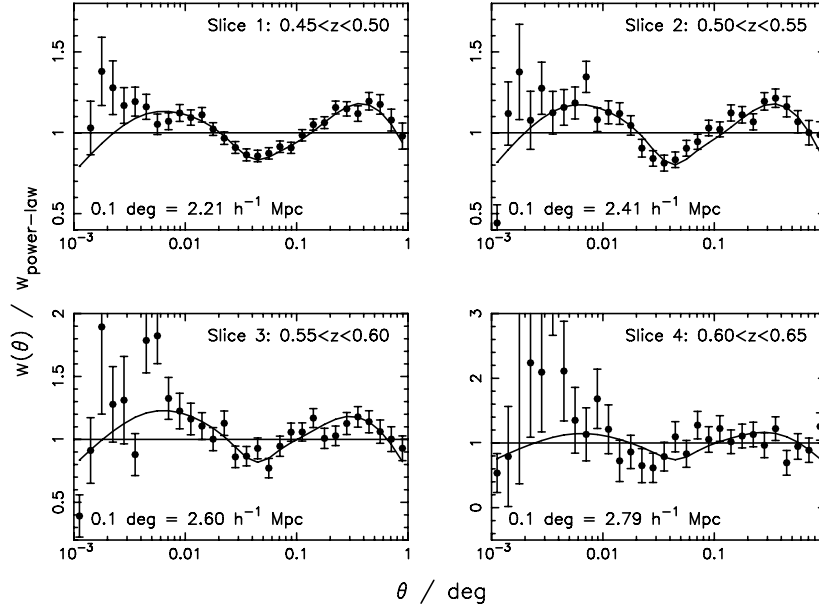


Figure 6. The measured angular correlation functions in the four redshift slices together with the best-fitting halo models. The y-axis is the ratio of the angular correlation function $w(\theta)$ and the best-fitting power-law model in each slice (from Section 3.3). The comoving spatial scale corresponding to an angular separation of 0.1 at the median redshift of each slice is displayed in the bottom left-hand corner of each plot. A flat Λ cold dark matter universe is assumed with cosmological parameters $\Omega_m = 0.3$, $h = 0.7$, $\Omega_b/\Omega_m = 0.15$, $\sigma_8 = 0.8$ and $n_s = 1$.

dark matter haloes and are highly biased tracers of the clustering pattern. Our values for the effective mass are in the same range as the LRG halo mass measured by Mandelbaum et al. (2006b) using galaxy–galaxy lensing ($M = 10^{13.83} h^{-1} M_\odot$ for the bright sample of Mandelbaum et al.). We note that the systematic increase in galaxy bias in each redshift slice is driven more by the increasing luminosity threshold rather than redshift evolution.

Fig. 7 plots the statistical range of allowed halo occupation distributions $\langle N|M \rangle$ in each of the four redshift slices assuming the parametric description of equations (26) and (27). The parameters σ_{cut} , M_0 and β were varied over a grid and the probability determined at each grid point using the χ^2 statistic. This probability distribution was used to construct the mean and standard deviation of the value of $\langle N|M \rangle$ as a function of halo mass M . For example, the average number of our galaxy sample hosted by a halo of mass $M = 10^{14.5} h^{-1} M_\odot$ is (5.5, 4.1, 2.6, 1.4) in the four redshift slices, decreasing systematically with redshift as the threshold luminosity increases. Broadly speaking, the effect of increasing luminosity is to shift the HOD to higher masses without significantly changing its shape, that is, shifting to the right-hand side in Fig. 7 in a similar fashion for central and satellite galaxies. If we weight the HOD by

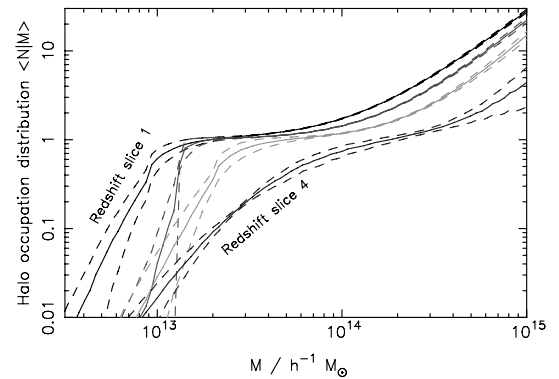


Figure 7. The range of halo occupation distributions fitted to each of the four redshift slices (with each subsequent slice shifting from left to right in the figure) assuming the parametric description of equations (26) and (27). This figure was generated by varying the parameters σ_{cut} , M_0 and β over a grid and determining the relative probability at each grid point using the χ^2 statistic. At each halo mass M , the probability distribution within the model parameter space was used to construct the mean HOD (the solid line) and the 68 per cent confidence region (the dotted lines).

the mass function of haloes which steeply decreases with increasing mass, we find that the fraction of galaxies that are classified as central galaxies is very high: $f_c = (0.88, 0.90, 0.93, 0.97)$ in the four redshift slices. This agrees with the standard picture of LRGs forming at the heart of massive dark matter haloes.

The best-fitting value of β , the slope of the power-law HOD for satellite galaxies, increases systematically with redshift. The most important cause of this trend is evolution of β with luminosity, as measured at low redshift (Zehavi et al. 2005b). Investigating this further, we refitted the halo-model parameters in each of the redshift slices for each of the absolute i -band magnitude thresholds $M_i - 5 \log_{10} h = (-22.23, -22.56, -22.87, -23.20)$ which correspond to the luminosity thresholds of the four redshift slices. Hence, all redshift slices can provide a sample with $M_i - 5 \log_{10} h < -23.20$, but only the first redshift slice contributes a sample with $M_i - 5 \log_{10} h < -22.23$. We obtained the probability distribution for β for each halo-model fit, marginalizing over the other model parameters, and combined the results for matched luminosity samples in different redshift slices. We find that the measurements of β at a fixed luminosity threshold agree well between the redshift slices, and the combined result for the four luminosity thresholds listed above is $\beta = (1.57 \pm 0.02, 1.68 \pm 0.03, 1.91 \pm 0.05, 2.15 \pm 0.14)$, confirming a significant evolution of β with luminosity. Considering just the first redshift slice, the best-fitting effective halo mass for the four luminosity slices is $\log_{10}(M_{\text{eff}}/h^{-1} M_\odot) = (13.61, 13.73, 13.88, 14.02)$ corresponding to number densities $n_g = (5.03, 2.45, 1.02, 0.32) \times 10^{-4} h^3 \text{ Mpc}^{-3}$.

Our results have a significant dependence on the assumed value of $\sigma_8 = 0.8$, which sets the overall normalization of the matter power spectrum. For $\sigma_8 = 0.7$, the best-fitting effective halo mass in each of the four redshift slices is $\log_{10}(M_{\text{eff}}/h^{-1} M_\odot) = (13.48, 13.54, 13.61, 13.68)$. For $\sigma_8 = 0.9$, the results are (13.72, 13.79, 13.85, 13.95). In all cases the minimum values of χ^2 are similar to those obtained for our default value of σ_8 , suggesting that there is a strong degeneracy between σ_8 and the halo-model parameters used in this analysis. The changing σ_8 affected the best-fitting value of M_0 much more strongly than the value of β . We also tried lowering the value of the scalar spectral index of the primordial power spectrum from $n_s = 1$ to 0.95, as supported by recent observations of the CMB radiation (Spergel et al. 2007). Assuming $\sigma_8 = 0.8$, the best-fitting effective masses were (13.59, 13.66, 13.72, 13.78) which do not differ significantly from our default values presented in Table 2.

6 COMPARISON TO PREVIOUS WORK

Several previous studies have fitted halo-model parameters to populations of red galaxies, for example, Magliocchetti & Porciani (2003, 2dFGRS), Zehavi et al. (2004, 2005b, SDSS), Collister & Lahav (2005, 2dFGRS groups catalogue), Phleps et al. (2006, COMBO-17 survey) and White et al. (2007, NDWFS). We make comparisons to these analyses below, where possible.

Halo-model fits to the 2dFGRS galaxy correlation function for late-type and early-type galaxies were performed by Magliocchetti & Porciani (2003). In addition, Collister & Lahav (2005) directly investigated the distribution of galaxies within 2dFGRS groups. These two studies produced a reasonably consistent measurement of the slope $\beta \approx 1$ of the HOD at high masses for early-type galaxies. Our best-fitting slope, $\beta = 1.5\text{--}2.0$, is much higher due to the following two factors. (1) The significantly higher luminosity of our galaxy samples; and (2) we make the distinction between central

and satellite galaxies, separating out a central galaxy contribution $N_c \approx 1$ at high masses. The latter has the effect of significantly steepening the slope of the power-law HOD fitted to the remaining satellite galaxies (which in fact contribute only 5–10 per cent of our sample, as noted in Section 5). In other words, we effectively fit a model $N = N_c (1 + N_s) = 1 + (M/M_0)^\beta$ at high masses, rather than $N = (M/M_0)^\beta$.

Zehavi et al. (2004) analysed a luminous subset of galaxies from the SDSS ‘main’ spectroscopic data base with $M_r < -21$ and mean redshift $z \approx 0.1$. They found that an HOD of the form

$$\begin{aligned} \langle N | M \rangle &= 0 \quad (M < M_{\text{cut}}) \\ &= 1 \quad (M_{\text{cut}} < M < M_0) \\ &= (M/M_0)^\beta \quad (M > M_0) \end{aligned} \quad (35)$$

produced a good fit to the clustering data, where $M_{\text{cut}} = 10^{12.79} h^{-1} M_\odot$, $M_0 = 10^{13.68} h^{-1} M_\odot$ and $\beta = 0.89$. The effective mass corresponding to these parameters is $M_{\text{eff}} = 10^{13.83} h^{-1} M_\odot$ (for their choice of $\sigma_8 = 0.9$). The number density of the Zehavi et al. (2004) sample is $n_g = 9.9 \times 10^{-4} h^3 \text{ Mpc}^{-3}$, which is a factor of 2–3 higher than our study. Although the effective masses are similar, we note that the redshift difference between the Zehavi et al. (2004) sample and ours may be important; the number density of a sample of dark matter haloes of fixed mass increases with decreasing redshift owing to the growth of structure. The difference in the best-fitting value of the power-law slope between Zehavi et al. and our analysis is again connected to the different forms of HOD fitted (owing to our central galaxy contribution, as discussed above).

Zehavi et al. (2005b) presented an extended analysis of the SDSS data in which central and satellite galaxy contributions are considered separately. Their default model includes a sharp cut-off for the central galaxy HOD at a fixed mass, rather than our ‘softened’ transition from 0 to 1 galaxies. They find that the slope of the power-law satellite HOD increases systematically with luminosity in a manner entirely consistent with our high-luminosity measurements of $\beta = 1.5\text{--}2.0$. In addition, Zehavi et al. (2005b) note that the step function for $\langle N_c | M \rangle$ produces a poor fit to the data in their highest-luminosity bin, consistent with our requirement for a softened transition parametrized by σ_{cut} . They also find, in agreement with our analysis, that the great majority of luminous galaxies are central galaxies of their host dark matter haloes, rather than satellites in more massive systems. A low ($\lesssim 10$ per cent) satellite fraction for the most-luminous elliptical galaxies is also found in galaxy–galaxy lensing studies (Seljak et al. 2005; Mandelbaum et al. 2006a) and other clustering studies (Tinker et al. 2007; van den Bosch et al. 2007).

Phleps et al. studied various populations of galaxies in the COMBO-17 survey at a mean redshift of $\bar{z} = 0.6$ which is similar to our data set. For red-sequence galaxies, Phleps et al. quote an effective halo mass for their best-fitting model of $M_{\text{eff}} = 10^{13.2} h^{-1} M_\odot$, whereas we find $M_{\text{eff}} = 10^{13.7} h^{-1} M_\odot$ (Table 2). This apparently large discrepancy is caused by the significant difference in the luminosity threshold of the two samples: the number density of our LRG catalogue is more than an order of magnitude smaller (there is also a difference in the assumed value of σ_8).

White et al. (2007) fitted an HOD model to the clustering of LRGs in the NOAO Deep Wide-Field Survey (NDWFS) Bootes field of 9 deg^2 , analysed in redshift slices between $z = 0.4$ and 1.0 . The luminosity thresholds are fixed such that the galaxy number density in each redshift slice is $10^{-3} h^3 \text{ Mpc}^{-3}$, exceeding our sample by a factor of ≈ 3 at $z = 0.5$. White et al. demonstrated that the clustering of the $z = 0.5$ sample cannot be accounted for by simple passive evolution of the $z = 0.9$ sample, but rather there must be merging or

disruption of the most-luminous satellite galaxies in massive haloes. The best-fitting satellite fraction in the NDWFS sample is found to be 18 per cent, a little higher than the results of our study, but consistent with a trend in which satellite fraction decreases with increasing luminosity.

In conclusion, our halo-model parameter measurements appear broadly consistent with previous work, allowing for differing luminosity thresholds. A fully consistent comparison of our analysis at $z \approx 0.55$ with results at $z \approx 0$ is beyond the scope of this work, owing to the differing forms of HOD assumed by different authors, but a topic worthy of further investigation.

Measurements of the three-point clustering functions will add further insight into the LRG clustering properties. Recent work by Kulkarni et al. (2007), analysing the SDSS spectroscopic LRG sample at $z \approx 0.35$, favoured a shallower slope for the satellite HOD, $\beta \approx 1.4$, with a higher satellite fraction of 17 per cent. Further study is required to understand these differences.

7 TESTS FOR SYSTEMATIC PHOTOMETRIC ERRORS

We performed a series of tests for potential systematic photometric errors that may affect our clustering results. Following the discussion in Blake et al. (2007), we compared the angular correlation function measured for the ‘default’ sample with that obtained by restricting or extending the galaxy selection in the following ways.

- (i) Exclusion of areas of high dust extinction (>0.1 mag);
- (ii) exclusion of areas of poor astronomical seeing (>1.5 arcsec);
- (iii) exclusion of areas lying in the overlap regions between survey stripes.
- (iv) Exclusion of areas in the vicinity of very bright objects (circular masks of radius 1 arcmin around objects with $i < 12$); and
- (v) variations in the star–galaxy separation criteria. This is quantified by the coefficient δ_{sg} in Blake et al. (2007) and Collister et al. (2007), which fixes the aggressiveness of the star–galaxy separation

in the neural network. Our default choice is $\delta_{\text{sg}} > 0.2$, which results in a level of stellar contamination of 1.5 per cent with the loss of only ~ 0.1 per cent of the genuine galaxies (see fig. 13 in Collister et al. 2007). We also tried $\delta_{\text{sg}} > 0$ (no additional star–galaxy separation in the neural network; stellar contamination 4.4 per cent) and $\delta_{\text{sg}} > 0.8$ (stellar contamination 0.4 per cent; loss of 1.2 per cent genuine galaxies).

We refer the reader to Blake et al. (2007) for a more thorough discussion of the possible effects of these systematic errors.

Our results are presented in Figs 8–10. Each plot is composed of four panels, one for each redshift slice. In each panel we show the angular correlation function for the default sample together with that corresponding to a change in the galaxy selection criteria. We divide all the correlation functions by a power-law fit to the default measurement to render the results more easily comparable.

We conclude from Figs 8 and 9 that our results are robust against the details of the angular selection function: varying dust extinction, seeing, overlap regions and bright object masks all have little effect on the measured correlation function. Fig. 10 reveals that the details of the star–galaxy separation affect the amplitude of the measured correlation function although not (to first order) the shape. This amplitude shift is already encoded in the stellar contamination factor $(1 - f)^2$. In no case does a change in the galaxy selection alter the detectability or shape of the halo-model signature.

8 CONCLUSIONS

We have measured the angular correlation function of LRGs in the SDSS imaging survey, using accurate photometric redshifts to divide the galaxies into narrow redshift slices and create volume-limited samples. We find the following.

- (i) A canonical power-law fit provides a poor description of the small-scale angular correlation function, although the best-fitting slope $w(\theta) \propto \theta^{-0.95}$ agrees well with previous studies of LRGs.

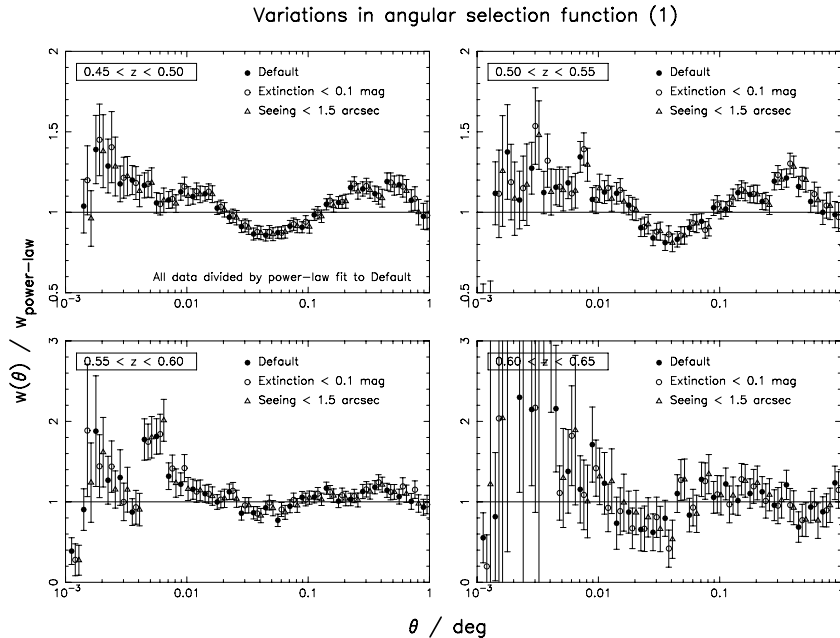


Figure 8. The dependence of the angular correlation function measurement in four redshift slices on varying *dust extinction* and *astronomical seeing*. Results from the default catalogue are compared to an analysis restricting the regions analysed to (1) a maximum dust extinction of 0.1 mag, or (2) a maximum seeing of 1.5 arcsec. The y-axis is the angular correlation function $w(\theta)$ divided by a power-law fit to the default model.

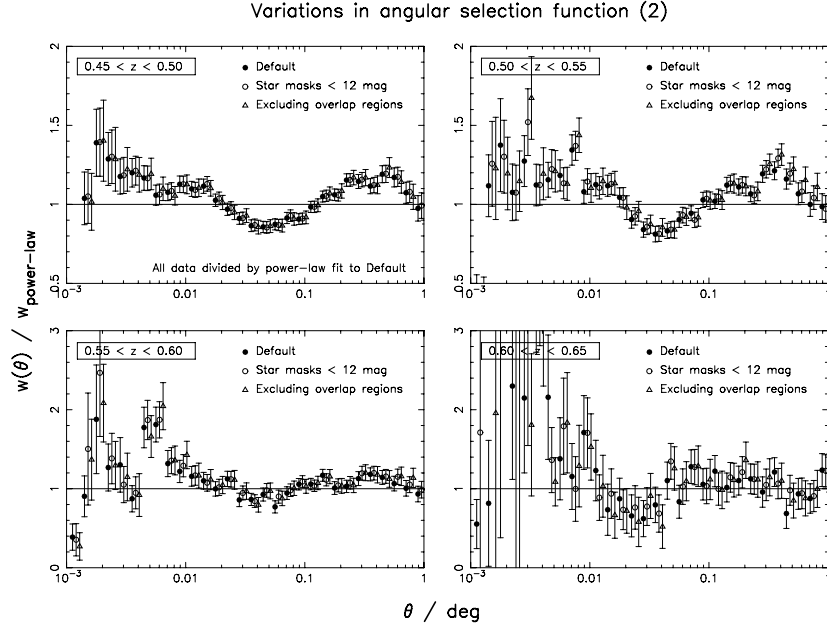


Figure 9. The dependence of the angular correlation function measurement in four redshift slices in the presence of *bright objects* and *stripe overlap regions*. Results from the default catalogue are compared to (1) an analysis placing circular masks of radius 1 arcmin around all objects with *i*-band magnitudes brighter than 12, and (2) an analysis excluding overlap regions between stripes. The results are displayed in the same manner as Fig. 8.

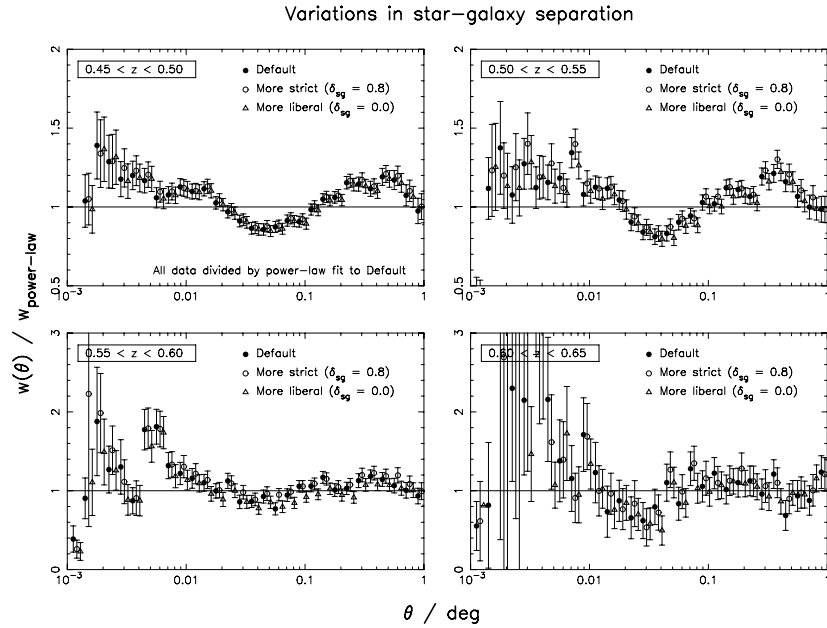


Figure 10. The dependence of the angular correlation function measurement in four redshift slices on the star–galaxy separation criteria. Results from the default catalogue are compared to analyses changing the value of the coefficient δ_{sg} defined in Blake et al. (2007) and Collister et al. (2007), which controls the aggressiveness of the star–galaxy separation. The results are displayed in the same manner as Fig. 8.

(ii) The halo model of galaxy clustering, composed of separate one-halo and two-halo contributions, produces a good fit to the deviations from a power law. We assume an HOD with separate components for central and satellite galaxies, implementing realistic models for scale-dependent bias, halo exclusion and non-linear growth of structure. We find that the HOD for central galaxies requires a ‘soft’ transition from zero to one central galaxy, as opposed to a step function, to reproduce the observations. The functional form $N_c = 0.5 \{1 + \text{erf}[\log_{10}(M/M_{\text{cut}})/\sigma_{\text{cut}}]\}$ provides a good fit for central galaxies, combined with a power-law HOD for satellite galaxies

$N_s = (M/M_0)^\beta$. One parameter of the model (M_{cut}) is fixed by the overall number density of the galaxy sample; hence, this halo model contains three variable parameters (σ_{cut} , M_0 and β), just one more than a simple power law. Allowing the concentration parameter c to vary, or including a more sophisticated HOD for satellite galaxies, does not improve the model fits.

(iii) The slope β of the power-law HOD for satellite galaxies is a strong function of luminosity, increasing to $\beta \approx 2$ for our most-luminous sample. This is consistent with extrapolating the variation of β with luminosity in local samples (Zehavi et al. 2005b). The

physical implication of this result is that haloes of higher mass have greater relative efficiency at producing high-luminosity satellites. We find no variation of β with redshift across our sample (from $z = 0.45$ to 0.65).

(iv) The best-fitting width σ_{cut} of the transition from zero to one central galaxy with increasing mass is in the range $\sigma_{\text{cut}} = 0.1$ – 0.5 (this is the parameter least well constrained by our analysis). The physical implication of this result is a broad scatter between central galaxy luminosity and host halo mass. This scatter results from galaxy formation physics and from the photo- z error in the conversion of apparent magnitude to luminosity.

(v) The halo-model fits describe how LRGs populate dark matter haloes as a function of their mass M . The average number of our galaxy sample hosted by a halo of mass $M = 10^{14.5} h^{-1} M_{\odot}$ is (5.5, 4.1, 2.6, 1.4) in the four redshift slices, decreasing systematically with redshift as the threshold luminosity increases. Broadly speaking, the effect of increasing luminosity is to shift the HOD uniformly to higher masses without significantly changing its shape. The large majority of galaxies in our sample are classified as central galaxies of their host dark matter haloes, rather than satellites in more massive systems, in agreement with previous studies of galaxy–galaxy lensing and clustering of the most-luminous galaxies. The satellite fraction varies in the range 3–12 per cent across the redshift slices.

(vi) The halo-model fits provide robust predictions of the average linear bias of the LRGs on large scales and the effective mass of their host dark matter haloes. The resulting amplitude of the linear bias ($b_g = 1.9$ – 2.6 , increasing with redshift, assuming a normalization of the matter power spectrum $\sigma_8 = 0.8$) agrees well with fits to the large-scale power spectrum (Blake et al. 2007). The effective halo mass ($M_{\text{eff}} = 10^{13.7}$ – $10^{13.8} h^{-1} M_{\odot}$) provides a quantitative statement of how the LRGs trace the underlying dark matter haloes. The value of M_{eff} has a significant dependence on σ_8 , and a weak dependence on the slope of the primordial scalar index n_s : the effective mass increases by $\Delta \log_{10}(M_{\text{eff}}/h^{-1} M_{\odot}) \approx 0.1$ when the value of σ_8 is increased from 0.8 to 0.9, and decreases by $\Delta \log_{10}(M_{\text{eff}}/h^{-1} M_{\odot}) \approx 0.02$ when n_s is decreased from 1.0 to 0.95.

Future studies will explore joint fits of the cosmological parameters and halo-model parameters (Abazajian et al. 2005; Zheng & Weinberg 2007), direct measurement of halo occupation via a cluster and group-finding analysis of the photo- z catalogue, a consistent comparison of clustering of LRGs at $z \approx 0.5$ and at $z \approx 0$, and testing the halo model further via three-point clustering statistics and higher moments.

ACKNOWLEDGMENTS

We are very grateful to an anonymous referee for extremely detailed comments that greatly improved this paper. We also thank Zheng Zheng for providing very helpful advice and code comparisons for the halo occupation distribution modelling. We thank Sarah Bridle, Filipe Abdalla, Ravi Sheth and Michael Brown for useful discussions. CB acknowledges support from the Izaak Walton Killam Memorial Fund for Advanced Studies and from the Canadian Institute for Theoretical Astrophysics National Fellowship programme. OL acknowledges a PPARC Senior Research Fellowship.

Funding for the SDSS has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

REFERENCES

- Abazajian K. et al., 2005, *ApJ*, 625, 613
- Amendola L., Quercellini C., Giallongo E., 2005, *MNRAS*, 357, 429
- Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, *MNRAS*, 311, 793
- Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
- Berlind A. A. et al., 2003, *ApJ*, 593, 1
- Blake C. A., Bridle S. L., 2005, *MNRAS*, 363, 1329
- Blake C. A., Collister A., Bridle S. L., Lahav O., 2007, *MNRAS*, 374, 1527
- Brown M. J. I., Dey A., Jannuzi B. T., Lauer T. R., Tiede G. P., Mikles V. J., 2003, *ApJ*, 597, 225
- Budavari T. et al., 2003, *ApJ*, 595, 59
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559
- Cannon R. et al., 2006, *MNRAS*, 372, 425
- Collister A., Lahav O., 2004, *PASP*, 116, 345
- Collister A., Lahav O., 2005, *MNRAS*, 361, 415
- Collister A. et al., 2007, *MNRAS*, 375, 68
- Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1
- Dolney D., Jain B., Takada M., 2004, *MNRAS*, 352, 1019
- Eisenstein D. J., Hu W., 1998, *ApJ*, 518, 2
- Eisenstein D. J. et al., 2001, *AJ*, 122, 2267
- Eisenstein D. J. et al., 2005a, *ApJ*, 619, 178
- Eisenstein D. J. et al., 2005b, *ApJ*, 633, 560
- Firth A. E., Lahav O., Somerville R. S., 2003, *MNRAS*, 339, 1195
- Fry J. N., 1996, *ApJ*, 461, 65
- Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, *MNRAS*, 321, 372
- Hawkins E. et al., 2003, *MNRAS*, 346, 78
- Hogg D. W. et al., 2003, *ApJ*, 585, L5
- Kauffmann G., Nusser A., Steinmetz M., 1997, *MNRAS*, 286, 795
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottloeber S., Allgood B., Primack J. R., 2004, *ApJ*, 609, 35
- Kulkarni G. V., Nichol R. C., Sheth R. K., Seo H.-J., Eisenstein D. J., Gray A., 2007, *MNRAS*, 378, 1196
- Lahav O. et al., 2002, *MNRAS*, 333, 961
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Madgwick D. S. et al., 2003, *MNRAS*, 344, 847
- Magliocchetti M., Porciani C., 2003, *MNRAS*, 346, 186
- Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006a, *MNRAS*, 368, 715
- Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006b, *MNRAS*, 372, 758
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Norberg P. et al., 2002, *MNRAS*, 332, 827
- Padmanabhan N. et al., 2005, *MNRAS*, 359, 237
- Padmanabhan N. et al., 2007, *MNRAS*, 378, 852
- Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144
- Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton
- Phleps S., Peacock J. A., Meisenheimer K., Wolf C., 2006, *A&A*, 457, 145
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425

- Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, *ApJ*, 546, 20
- Seljak U., 2000, *MNRAS*, 318, 203
- Seljak U. et al., 2005, *Phys. Rev. D*, 71, 43511
- Seo H., Eisenstein D. J., 2003, *ApJ*, 598, 720
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Smith R. E. et al., 2003, *MNRAS*, 341, 1311
- Spergel D. N. et al. 2007, *ApJS*, 170, 377
- Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, *ApJ*, 631, 41
- Tinker J. L., Norberg P., Weinberg D. H., Warren M. S. 2007, *ApJ*, 659, 877
- van den Bosch F. C., Yang X., Mo H. J., Weinmann S. M., Maccio A., More S., Cacciato M., Skibba R., Xi K., 2007, *MNRAS*, 376, 841
- Wake D. A. et al., 2006, *MNRAS*, 372, 537
- White M., Zheng Z., Brown M. J. I., Dey A., Jannuzi B. T., 2007, *ApJ*, 655, L69
- York D. G. et al., 2000, *AJ*, 120, 1579
- Zehavi I. et al., 2004, *ApJ*, 608, 16
- Zehavi I. et al., 2005a, *ApJ*, 621, 22
- Zehavi I. et al., 2005b, *ApJ*, 630, 1
- Zhan H., Knox L., Tyson J. A., Margoniner V., 2006, *ApJ*, 640, 8
- Zheng Z., 2004, *ApJ*, 610, 61
- Zheng Z., Weinberg D. H., 2007, *ApJ*, 659, 1
- Zheng Z. et al., 2005, *ApJ*, 633, 791

This paper has been typeset from a \LaTeX file prepared by the author.